

- 67 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (Other than US) AMRAD OPERATIONS PTY LTD
(US only) Douglas James HILTON, Nicos Antony NICOLA, Alison FARLEY, Tracey WILLSON, Jian-Guo ZHANG, Warren ALEXANDER, Steven RAKAR, Louis FABRI, Tetsuo KOJIMA, Masatsugu MAEDA, Yasumfumi KIKUCHI, Andrew NASH

(ii) TITLE OF INVENTION: A NOVEL HAEMPOIETIN RECEPTOR AND GENETIC SEQUENCES ENCODING SAME

(iii) NUMBER OF SEQUENCES: 54

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSOR
- (B) STREET: 400 GARDEN CITY PLAZA
- (C) CITY: GARDEN CITY
- (D) STATE: NEW YORK
- (E) COUNTRY: UNITED STATES OF AMERICA
- (F) ZIP: 11530-0299

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: Patent In Release #1.0, Version #1.25

(vii) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: CIP APPLICATION OF USSN 08/928,720
- (B) FILING DATE: 10-MAR-1998

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/GB97/02479
- (B) FILING DATE: 11-SEP-1997

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US APPLICATION NO. 08/928,720
- (B) FILING DATE: 11-SEP-1997

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: PO2246/96
- (B) FILING DATE: 11-SEP-1996

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: DIGIGLIO, FRANK S
- (B) REGISTRATION NO: 31,346
- (C) REFERENCE/DOCKET NUMBER: 10857Z

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: +516 742 4343
- (B) TELEFAX: +516 742 4366

00037657-031098

Sub
B

- 68 -

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Trp Ser Xaa Trp Ser

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACTCGCTCCA GATTCCCGCC TTTT

24

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TCCCGCCTTT TTCGACCCAT AGAT

24

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGTACTTGGC TTGGAAGAGG AAAT

24

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

09037657 "031098

- 69 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGGCTCACGT GCACGTCGGG TGGG

24

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCTGCTGTT AAAGGGCTTC TC

22

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

(A/G)CTCCA(A/G)TC(A/G) CTCCA

15

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

(A/G)CTCCA(C/T)TC(A/G) CTCCA

15

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

09037657-031096

- 70 -

AAGTGTGACC ATCATGTGGA C

21

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGAGGTGTTA AGGAGGCG

18

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGCCCCGCGG GTCGCCCCG

18

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1506 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCACGAGCT TCGCTGTCCG CGCCCACTGA CGCGCGTGCG GACCCGAGCC CCAATCTGCA

-64

CCCCGCAGAC TCGCCCCCGC CCCATACCGG CGTTGCAGTC ACCGCCCGTT GCGCGCCACC

-4

CCC

-3

ATG CCC GCG GGT CGC CCG GGC CCC GTC GCC CAA TCC GCG CGG CGG CCG
 Met Pro Ala Gly Arg Pro Gly Pro Val Ala Gln Ser Ala Arg Arg Pro
 1 5 10 15

48

CCG CGG CCG CTG TCC TCG CTG TGG TCG CCT CTG TTG CTC TGT GTC CTC
 Pro Arg Pro Leu Ser S r Leu Trp Ser Pro Leu Leu Leu Cys Val Leu
 20 25 30

96

05037657-031099B

- 71 -

GGG GTG CCT CGG GGC GGA TCG GGA GCC CAC ACA GCT GTA ATC AGC CCC Gly Val Pro Arg Gly Gly Ser Gly Ala His Thr Ala Val Ile Ser Pro 35 40 45	144
CAG GAC CCC ACC CTT CTC ATC GGC TCC TCC CTG CAA GCT ACC TGC TCT Gln Asp Pro Thr Leu Leu Ile Gly Ser Ser Leu Gln Ala Thr Cys Ser 50 55 60	192
ATA CAT GGA GAC ACA CCT GGG GCC ACC GCT GAG GGG CTC TAC TGG ACC Ile His Gly Asp Thr Pro Gly Ala Thr Ala Glu Gly Leu Tyr Trp Thr 65 70 75 80	240
CTC AAT GGT CGC CGC CTG CCC TCT GAG CTG TCC CGC CTC CTT AAC ACC Leu Asn Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr 85 90 95	288
TCC ACC CTG GCC CTG GCC CTG GCT AAC CTT AAT GGG TCC AGG CAG CAG Ser Thr Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln 100 105 110	336
TCA GGA GAC AAT CTG GTG TGT CAC GCC CGA GAC GGC AGC ATT CTG GCT Ser Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu Ala 115 120 125	384
GGC TCC TGC CTC TAT GTT GGC TTG CCC CCT GAG AAG CCC TTT AAC ATC Gly Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Phe Asn Ile 130 135 140	432
ACC TGC TGG TCC CGG AAC ATG AAG GAT CTC ACG TGC CGC TGG ACA CCG Ser Cys Trp Ser Arg Asn Met Lys Asp Leu Thr Cys Arg Trp Thr Pro 145 150 155 160	480
GGT GCA CAC GGG GAG ACA TTC TTA CAT ACC AAC TAC TCC CTC AAG TAC Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr 165 170 175	528
AAG CTG AGG TGG TAC GGT CAG GAT AAC ACA TGT GAG GAG TAC CAC ACT Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His Thr 180 185 190	576
GTG GGC CCT CAC TCA TGC CAT ATC CCC AAG GAC CTG GCC CTC TTC ACT Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe Thr 195 200 205	624
CCC TAT GAG ATC TGG GTG GAA GCC ACC AAT CGC CTA GGC TCA GCA AGA Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala Arg 210 215 220	672
TCT GAT GTC CTC ACA CTG GAT GTC CTG GAC GTG GTG ACC ACG GAC CCC Ser Asp Val Leu Thr Leu Asp Val Leu Asp Val Val Thr Thr Asp Pro 225 230 235 240	720
CCA CCC GAC GTG CAC GTG AGC CGC GTT GGG GGC CTG GAG GAC CAG CTG Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln Leu 245 250 255	768
AGT GTG CGC TGG GTC TCA CCA CCA GCT CTC AAG GAT TTC CTC TTC CAA Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe Gln 260 265 270	816
GCC AAG TAC CAG ATC CGC TAC CGC GTG GAG GAC AGC GTG GAC TGG AAG Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp Lys 275 280 285	864
GTG GTG GAT GAC GTC AGC AAC CAG ACC TCC TGC CGT CTC GCG GGC CTG Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly Leu 290 295 300	912
AAG CCC GGC ACC GTT TAC TTC GTC CAA GTG CGT TGT AAC CCA TTC GGG Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly 305 310 315 320	960
ATC TAT GGG TCG AAA AAG GCG GGA ATC TGG AGC GAG TGG AGC CAC CCC Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro 325 330 335 340	1008

00037657-031000

- 72 -

325	330	335	
ACC GCT GCC TCC ACC CCT CGA AGT GAG CGC CCG GGC CCG GGC GGC GGG Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Gly 340 345 350			1056
GTG TGC GAG CCG CGG GGC GGC GAG CCC AGC TCG GGC CCG GTG CCG CGC Val Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg 355 360 365			1104
GAG CTC AAG CAG TTC CTC GGC TGG CTC AAG AAG CAC GCA TAC TGC TCG Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser 370 375 380			1152
AAC CTT AGT TTC CGC CTG TAC GAC CAG TGG CGT GCT TGG ATG CAG AAG Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys 385 390 395 400			1200
TCA CAC AAG ACC CGA AAC CAG GTC CTG CCG GCT AAA CTC TAAGGATAGG Ser His Lys Thr Arg Asn Gln Val Leu Pro Ala Lys Leu 405 410			1249
CCATCCTCCT GCTGGGTGAG ACCTGGAGGC TCACCTGAAT TGGAGCCCCCT CTGTACCATC			1309
TGGGCAACAA AGAAACCTAC CAGAGGCTGG GGCACAATGA GCTCCACAA CCACAGCTTT			1369
GGTCCACATG ATGGTCACAC TTGGATATAC CCCAGTGTGG GTAAGGTTGG GGTATTGCAG			1429
GGCCTCCCAA CAATCTCTTT AAATAAATAA AGGAGTTGTT CAGGTAAAAA AAAAAAAAAA			1489
AAAAAAAAAA AAAAAA			1506

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 413 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Pro	Ala	Gly	Arg	Pro	Gly	Pro	Val	Ala	Gln	Ser	Ala	Arg	Arg	Pro
1				5					10					15	
Pro	Arg	Pro	Leu	Ser	Ser	Leu	Trp	Ser	Pro	Leu	Leu	Leu	Cys	Val	Leu
			20					25					30		
Gly	Val	Pro	Arg	Gly	Gly	Ser	Gly	Ala	His	Thr	Ala	Val	Ile	Ser	Pro
			35				40					45			
Gln	Asp	Pro	Thr	Leu	Leu	Ile	Gly	Ser	Ser	Leu	Gln	Ala	Thr	Cys	Ser
	50					55				60					
Ile	His	Gly	Asp	Thr	Pro	Gly	Ala	Thr	Ala	Glu	Gly	Leu	Tyr	Trp	Thr
	65				70				75					80	
Leu	Asn	Gly	Arg	Arg	Leu	Pro	Ser	Glu	Leu	Ser	Arg	Leu	Leu	Asn	Thr
				85					90					95	
Ser	Thr	Leu	Ala	Leu	Ala	Leu	Ala	Asn	Leu	Asn	Gly	Ser	Arg	Gln	Gln
			100					105					110		
Ser	Gly	Asp	Asn	Leu	Val	Cys	His	Ala	Arg	Asp	Gly	Ser	Ile	Leu	Ala
		115					120				125				
Gly	Ser	Cys	Leu	Tyr	Val	Gly	Leu	Pro	Pro	Glu	Lys	Pro	Phe	Asn	Ile
		130				135					140				
Ser	Cys	Trp	Ser	Arg	Asn	Met	Lys	Asp	Leu	Thr	Cys	Arg	Trp	Thr	Pro
					150					155					160

09037657-031098

- 73 -

Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr
 165 170 175
 Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His Thr
 180 185 190
 Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe Thr
 195 200 205
 Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala Arg
 210 215 220
 Ser Asp Val Leu Thr Leu Asp Val Leu Asp Val Thr Thr Asp Pro
 225 230 235 240
 Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln Leu
 245 250 255
 Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe Gln
 260 265 270
 Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp Lys
 275 280 285
 Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly Leu
 290 295 300
 Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly
 305 310 315 320
 Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro
 325 330 335
 Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Gly
 340 345 350
 Val Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg
 355 360 365
 Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser
 370 375 380
 Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys
 385 390 395 400
 Ser His Lys Thr Arg Asn Gln Val Leu Pro Ala Lys Leu
 405 410

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1549 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1278

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGCACGAGCT TCGCTGTCCG CGCCCACTGA CGCGCGTGG GACCCGAGCC CCAATCTGCA
 CCCCAGAC TCGCCCCCGC CCCATACCG CGTTGCAGTC ACCGCCCGTT GCGCGCCACC

-65

-5

 05037657-031099
 05037657-031099

- 74 -

CCCA

-1

ATG CCC GCG GGT CGC CCG GGC CCC GTC GCC CAA TCC GCG CGG CGG CCG	48
Met Pro Ala Gly Arg Pro Gly Pro Val Ala Gln Ser Ala Arg Arg Pro	
1 5 10 15	
CCG CGG CCG CTG TCC TCG CTG TGG TCG CCT CTG TTG CTC TGT GTC CTC	96
Pro Arg Pro Leu Ser Ser Leu Trp Ser Pro Leu Leu Leu Cys Val Leu	
20 25 30	
GGG GTG CCT CGG GGC GGA TCG GGA GCC CAC ACA GCT GTA ATC AGC CCC	144
Gly Val Pro Arg Gly Gly Ser Gly Ala His Thr Ala Val Ile Ser Pro	
35 40 45	
CAG GAC CCC ACC CTT CTC ATC GGC TCC TCC CTG CAA GCT ACC TGC TCT	192
Gln Asp Pro Thr Leu Leu Ile Gly Ser Ser Leu Gln Ala Thr Cys Ser	
50 55 60	
ATA CAT GGA GAC ACA CCT GGG GCC ACC GCT GAG GGG CTC TAC TGG ACC	240
Ile His Gly Asp Thr Pro Gly Ala Thr Ala Glu Gly Leu Tyr Trp Thr	
65 70 75 80	
CTC AAT GGT CGC CGC CTG CCC TCT GAG CTG TCC CGC CTC CTT AAC ACC	288
Leu Asn Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr	
85 90 95	
TCC ACC CTG GCC CTG GCC CTG GCT AAC CTT AAT GGG TCC AGG CAG CAG	336
Ser Thr Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln	
100 105 110	
TCA GGA GAC AAT CTG GTG TGT CAC GCC CGA GAC GGC AGC ATT CTG GCT	384
Ser Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu Ala	
115 120 125	
GGC TCC TGC CTC TAT GTT GGC TTG CCC CCT GAG AAG CCC TTT AAC ATC	432
Gly Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Phe Asn Ile	
130 135 140	
AGC TGC TGG TCC CGG AAC ATG AAG GAT CTC ACG TGC CGC TGG ACA CCG	480
Ser Cys Trp Ser Arg Asn Met Lys Asp Leu Thr Cys Arg Trp Thr Pro	
145 150 155 160	
GGT GCA CAC GGG GAG ACA TTC TTA CAT ACC AAC TAC TCC CTC AAG TAC	528
Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr	
165 170 175	
AAG CTG AGG TGG TAC GGT CAG GAT AAC ACA TGT GAG GAG TAC CAC ACT	576
Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His Thr	
180 185 190	
GTG GGC CCT CAC TCA TGC CAT ATC CCC AAG GAC CTG GCC CTC TTC ACT	624
Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe Thr	
195 200 205	
CCC TAT GAG ATC TGG GTG GAA GCC ACC AAT CGC CTA GGC TCA GCA AGA	672
Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala Arg	
210 215 220	
TCT GAT GTC CTC ACA CTG GAT GTC CTG GAC GTG GTG ACC ACG GAC CCC	720
Ser Asp Val Leu Thr Leu Asp Val Leu Asp Val Val Thr Thr Asp Pro	
225 230 235 240	
CCA CCC GAC GTG CAC GTG AGC CGC GTT GGG GGC CTG GAG GAC CAG CTG	768
Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln Leu	
245 250 255	
AGT GTG CGC TGG GTC TCA CCA CCA GCT CTC AAG GAT TTC CTC TTC CAA	816
Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe Gln	
260 265 270	
GCC AAG TAC CAG ATC CGC TAC CGC GTG GAG GAC AGC GTG GAC TGG AAG	864
Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp Lys	
275 280 285	

05037657.031098

- 75 -

GTG GTG GAT GAC GTC AGC AAC CAG ACC TCC TGC CGT CTC GCG GGC CTG 912
 Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly Leu
 290 295 300

AAG CCC GGC ACC GTT TAC TTC GTC CAA GTG CGT TGT AAC CCA TTC GGG 960
 Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly
 305 310 315 320

ATC TAT GGG TCG AAA AAG GCG GGA ATC TGG AGC GAG TGG AGC CAC CCC 1008
 Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro
 325 330 335

ACC GCT GCC TCC ACC CCT CGA AGT GAG CGC CCG GGC CCG GGC GGC GGG 1056
 Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Gly
 340 345 350

GTG TGC GAG CCG CGG GGC GGC GAG CCC AGC TCG GCC CCG GTG CGG CGC 1104
 Val Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg
 355 360 365

GAG CTC AAG CAG TTC CTC GGC TGG CTC AAG AAG CAC GCA TAC TGC TCG 1152
 Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser
 370 375 380

AAC CTT AGT TTC CGC CTG TAC GAC CAG TGG CGT GCT TGG ATG CAG AAG 1200
 Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys
 385 390 395 400

TCA CAC AAG ACC CGA AAC CAG GAC GAG GGG ATC CTG CCT TCG GGC AGA 1248
 Ser His Lys Thr Arg Asn Gln Asp Glu Gly Ile Leu Pro Ser Gly Arg
 405 410 415

CGG GGT GCG GCG AGA GGT CCT GCC GGT TAAACTCTAA GGATAGGCCA 1295
 Arg Gly Ala Ala Arg Gly Pro Ala Gly
 420 425

TCCTCTGCT GGGTCAGACC TGGAGGCTCA CCTGAATTGG AGCCCCCTCTG TACCATCTGG 1355

GCAACAAAGA AACCTACCAG AGGCTGGGGC ACAATGAGCT CCCACAACCA CAGCTTTGGT 1415

CCACATGATG GTCACACTTG GATATACCCC AGTGTTGGGTA AGGTTGGGGT ATTGCAGGGC 1475

CTCCTCAACAA TCTCTTTAAA TAAATAAAGG AGTTOTTCAG GTAAAAA AAAA 1535

AAAAAA AAAA 1549

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 425 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Pro Ala Gly Arg Pro Gly Pro Val Ala Gln Ser Ala Arg Arg Pro
 1 5 10 15

Pro Arg Pro Leu Ser Ser Leu Trp Ser Pro Leu Leu Leu Cys Val Leu
 20 25 30

Gly Val Pro Arg Gly Gly Ser Gly Ala His Thr Ala Val Ile Ser Pro
 35 40 45

Gln Asp Pro Thr Leu Leu Ile Gly Ser Ser Leu Gln Ala Thr Cys Ser
 50 55 60

Ile His Gly Asp Thr Pro Gly Ala Thr Ala Glu Gly Leu Tyr Trp Thr
 65 70 75 80

RECEIVED 1997-03-10

- 76 -

Leu Asn Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr
85 90 95

Ser Thr Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln
100 105 110

Ser Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu Ala
115 120 125

Gly Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Phe Asn Ile
130 135 140

Ser Cys Trp Ser Arg Asn Met Lys Asp Leu Thr Cys Arg Trp Thr Pro
145 150 155 160

Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr
165 170 175

Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His Thr
180 185 190

Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe Thr
195 200 205

Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala Arg
210 215 220

Ser Asp Val Leu Thr Leu Asp Val Leu Asp Val Val Thr Thr Asp Pro
225 230 235 240

Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln Leu
245 250 255

Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe Gln
260 265 270

Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp Lys
275 280 285

Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly Leu
290 295 300

Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly
305 310 315 320

Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro
325 330 335

Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Gly
340 345 350

Val Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg
355 360 365

Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser
370 375 380

Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys
385 390 395 400

Ser His Lys Thr Arg Asn Gln Asp Glu Gly Ile Leu Pro Ser Gly Arg
405 410 415

Arg Gly Ala Ala Arg Gly Pro Ala Gly
420 425

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 938 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

09037657-031098

- 77 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..468

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGC ACC GTT TAC TTC GTC CAA GTG CGT TGT AAC CCA TTC GGG ATC TAT	48
Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly Ile Tyr	
1 5 10 15	
GGG TCG AAA AAG GCG GGA ATC TGG AGC GAG TGG AGC CAC CCC ACC GCT	96
Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro Thr Ala	
20 25 30	
GCC TCC ACC CCT CGA AGT GAG CGC CCG GGC CCG GGC GGC GGG GTG TGC	144
Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Gly Val Cys	
35 40 45	
GAG CCG CGG GGC GGC GAG CCC AGC TCG GGC CCG GTG CGG CGC GAG CTC	192
Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg Glu Leu	
50 55 60	
AAG CAG TTC CTC GGC TGG CTC AAG AAG CAC GCA TAC TGC TCG AAC CTT	240
Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser Asn Leu	
65 70 75 80	
AGT TTC CGC CTG TAC GAC CAG TGG CGT GCT TGG ATG CAG AAG TCA CAC	288
Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys Ser His	
85 90 95	
AAG ACC CGA AAC CAG GTA GGA AAG TTG GGG GAG GCT TGC GTG GGG GGT	336
Lys Thr Arg Asn Gln Val Gly Lys Leu Gly Glu Ala Cys Val Gly Gly	
100 105 110	
AAA GGA GCA GAG GAA GAG AGA GAC CCG GGT GAG CAG CCT CCA CAA CAC	384
Lys Gly Ala Glu Glu Glu Arg Asp Pro Gly Glu Gln Pro Pro Gln His	
115 120 125	
CGC ACT CTT CTT TCC AAG CAC AGG ACG AGG GGA TCC TGC CCT CGG GCA	432
Arg Thr Leu Leu Ser Lys His Arg Thr Arg Gly Ser Cys Pro Arg Ala	
130 135 140	
GAC GGG GTG CGG CGA GAG GTA AGG GGG TCT GGG TGAGTGGGGC CTACAGCAGT	485
Asp Gly Val Arg Arg Glu Val Arg Gly Ser Gly	
145 150 155	
CTAGATGAGG CCCTTTCCCC TCCTTCGGTG TTGCTCAAAG GGATCTCTTA GTGCTCATTT	545
CACCCACTGC AAAGAGCCCC AGGTTTTACT GCATCATCAA GTTGCTGAAG GGTCCAGGCT	605
TAATGTGGCC TCTTTTCTGC CCTCAGGTCC TGCCGGCTAA ACTCTAAGGA TAGGCCATCC	665
TCCTGCTGGG TCAGACCTGG AGGCTCACCT GAATTGGAGC CCCTCTGTAC CTATCTGGGC	725
AACAAAGAAA CCTACCATGA GGCTGGGGCA CAATGAGCTC CCACAACCAC AGCTTTGGTC	785
CACATGATGG TCACACTTGG ATATACCCCA GTGTGGGTAA GGTGGGGTA TTGCAGGGCC	845
TCCCAACAAT CTCTTTAAAT AAATAAAGGA GTTGTTGAGG TAAAAA AAAA	905
AAAAA AAAA AAAA AAAA AAAA	938

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

09037657.031098

- 78 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly Ile Tyr
 1 5 10 15
 Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His Pro Thr Ala
 20 25 30
 Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly Val Cys
 35 40 45
 Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg Arg Glu Leu
 50 55 60
 Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys Ser Asn Leu
 65 70 75 80
 Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln Lys Ser His
 85 90 95
 Lys Thr Arg Asn Gln Val Gly Lys Leu Gly Glu Ala Cys Val Gly Gly
 100 105 110
 Lys Gly Ala Glu Glu Glu Arg Asp Pro Gly Glu Gln Pro Pro Gln His
 115 120 125
 Arg Thr Leu Leu Ser Lys His Arg Thr Arg Gly Ser Cys Pro Arg Ala
 130 135 140
 Asp Gly Val Arg Arg Glu Val Arg Gly Ser Gly
 145 150 155

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 834 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..834

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCC ACC CTT CTC ATC GGC TCC TCC CTG CAA GCT ACC TGC TCT ATA CAT	98
Pro Thr Leu Leu Ile Gly Ser Ser Leu Gln Ala Thr Cys Ser Ile His	
51 55 60 65	
GGA GAC ACA CCT GGG GCC ACC GCT GAG GGG CTC TAC TGG ACC CTC AAT	146
Gly Asp Thr Pro Gly Ala Thr Ala Glu Gly Leu Tyr Trp Thr Leu Asn	
70 75 80	
GGT CGC CGC CTG CCC TCT GAG CTG TCC CGC CTC CTT AAC ACC TCC ACC	194
Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr Ser Thr	
85 90 95	
CTG GCC CTG GCC CTG GCT AAC CTT AAT GGG TCC AGG CAG CAG TCA GGA	242
Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln Ser Gly	
100 105 110	
GAC AAT CTG GTG TGT CAC GCC CGA GAC GGC ACC ATT CTG GCT GGC TCC	290

09037657-031098

- 79 -

Asp	Asn	Leu	Val	Cys	His	Ala	Arg	Asp	Gly	Ser	Ile	Leu	Ala	Gly	Ser			
115					120					125					130			
TGC	CTC	TAT	GTT	GGC	TTG	CCC	CCT	GAG	AAG	CCC	TTT	AAC	ATC	AGC	TGC			338
Cys	Leu	Tyr	Val	Gly	Leu	Pro	Pro	Glu	Lys	Pro	Phe	Asn	Ile	Ser	Cys			
				135					140					145				
TGG	TCC	CGG	AAC	ATG	AAG	GAT	CTC	ACG	TGC	CGC	TGG	ACA	CCG	GGT	GCA			386
Trp	Ser	Arg	Asn	Met	Lys	Asp	Leu	Thr	Cys	Arg	Trp	Thr	Pro	Gly	Ala			
			150					155					200					
CAC	GGG	GAG	ACA	TTC	TTA	CAT	ACC	AAC	TAC	TCC	CTC	AAG	TAC	AAG	CTG			434
His	Gly	Glu	Thr	Phe	Leu	His	Thr	Asn	Tyr	Ser	Leu	Lys	Tyr	Lys	Leu			
			205				210					215						
AGG	TGG	TAC	GGT	CAG	GAT	AAC	ACA	TGT	GAG	GAG	TAC	CAC	ACT	GTG	GGG			482
Arg	Trp	Tyr	Gly	Gln	Asp	Asn	Thr	Cys	Glu	Glu	Tyr	His	Thr	Val	Gly			
		220			225						230							
CCC	CAC	TCA	TGC	CAT	ATC	CCC	AAG	GAC	CTG	GCC	CTC	TTC	ACT	CCC	TAT			530
Pro	His	Ser	Cys	His	Ile	Pro	Lys	Asp	Leu	Ala	Leu	Phe	Thr	Pro	Tyr			
					240					245					250			
GAG	ATC	TGG	GTG	GAA	GCC	ACC	AAT	CGC	CTA	GGC	TCA	GCA	AGA	TCT	GAT			578
Glu	Ile	Trp	Val	Glu	Ala	Thr	Asn	Arg	Leu	Gly	Ser	Ala	Arg	Ser	Asp			
				255					260					265				
GTC	CTC	ACA	CTG	GAT	GTC	CTG	GAC	GTG	GTG	ACC	ACG	GAC	CCC	CCA	CCC			626
Val	Leu	Thr	Leu	Asp	Val	Leu	Asp	Val	Val	Thr	Thr	Asp	Pro	Pro	Pro			
			270					275					280					
GAC	GTG	CAC	GTG	AGC	CGC	GTT	GGG	GGC	CTG	GAG	GAC	CAG	CTG	AGT	GTG			674
Asp	Val	His	Val	Ser	Arg	Val	Gly	Gly	Leu	Glu	Asp	Gln	Leu	Ser	Val			
			285				290					295						
CGC	TGG	GTC	TCA	CCA	CCA	GCT	CTC	AAG	GAT	TTC	CTC	TTC	CAA	GCC	AAG			722
Arg	Trp	Val	Ser	Pro	Pro	Ala	Leu	Lys	Asp	Phe	Leu	Phe	Gln	Ala	Lys			
			300			305					310							
TAC	CAG	ATC	CGC	TAC	CGC	GTG	GAG	GAC	AGC	GTG	GAC	TGG	AAG	GTG	GTG			770
Tyr	Gln	Ile	Arg	Tyr	Arg	Val	Glu	Asp	Ser	Val	Asp	Trp	Lys	Val	Val			
					320				325					330				
GAT	GAC	GTC	AGC	AAC	CAG	ACC	TCC	TGC	CGT	CTC	GCG	GGC	CTG	AAG	CCC			818
Asp	Asp	Val	Ser	Asn	Gln	Thr	Ser	Cys	Arg	Leu	Ala	Gly	Leu	Lys	Pro			
				335					340				345					
GGC	ACC	GTT	TAC	TTC	GTC	CAA	GTG	CGT	TGT	AAC	CCA	TTC	GGG	ATC	TAT			866
Gly	Thr	Val	Tyr	Phe	Val	Gln	Val	Arg	Cys	Asn	Pro	Phe	Gly	Ile	Tyr			
				350				355					360					
GGG	TCG	AAA	AAG	GCG	GGA													894
Gly	Ser	Lys	Lys	Ala	Gly													
				365														

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 278 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Pro	Thr	Leu	Leu	Ile	Gly	Ser	Ser	Leu	Gln	Ala	Thr	Cys	Ser	Ile	His		
51				55				60						65			
Gly	Asp	Thr	Pro	Gly	Ala	Thr	Ala	Glu	Gly	Leu	Tyr	Trp	Thr	Leu	Asn		
			70					75					80				

05037657-031098

- 80 -

Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr Ser Thr
85 90 95

Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln Ser Gly
100 105 110

Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu Ala Gly Ser
115 120 125 130

Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Phe Asn Ile Ser Cys
135 140 145

Trp Ser Arg Asn Met Lys Asp Leu Thr Cys Arg Trp Thr Pro Gly Ala
150 155 200

His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr Lys Leu
205 210 215

Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His Thr Val Gly
220 225 230

Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe Thr Pro Tyr
235 240 245 250

Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala Arg Ser Asp
255 260 265

Val Leu Thr Leu Asp Val Leu Asp Val Val Thr Thr Asp Pro Pro Pro
270 275 280

Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln Leu Ser Val
285 290 295

Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe Gln Ala Lys
300 305 310

Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp Lys Val Val
315 320 325 330

Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly Leu Lys Pro
335 340 345

Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe Gly Ile Tyr
350 355 360

Gly Ser Lys Lys Ala Gly
365

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 143 base pairs
 (B) TYPE: nucleic acids
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCATGAAGG CTTAGGGTGG GGATCGGTAG GACCCATGCA CCCAGAGAAA GGGACTGGTG	60
GCAACTTTCA AACTCTCTGG GGAAGGAAGA AGGGCTGAAA GAGG	104
ATG AAC GGG CTC AGA CAC AGC TGT AAT CAG CCC CCA GGA	143
Met Asn Gly Leu Arg His Ser Cys Asn Gln Pro Pro Gly	
5 10	

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 13 amino acids

00037657 "031098

- 81 -

(B) TYPE: amino acids
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Asn Gly Leu Arg His Ser Cys Asn Gln Pro Pro Gly
5 10

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1930 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCAGAGCT TCGCTGTCCG CGCCAGTGA CGCGGTGCG GACCCGAGCC CCAATCTGCA	60
CCCCGCAGAC TCGCCCCCGC CCCATACCGG CGTTGCAGTC ACCGCCCCGT GCGCGCCACC	120
CCCAATGCCC GCGGTCGCC CGGCCCCCGT CGCCCAATCC GCGCGCGGC CGCCGCGGCC	180
GCTGTCTCTG CTGTGGTCGC CTCTGTTGCT CTGTGTCCTC GGGGTGCCTC GGGGCGGATC	240
GGGAGCCCAC ACAGCTGTAA TCAGCCCCCA GGACCCACCC CTTCTCATCG GCTCCTCCCT	300
GCAAGCTACC TGCTCTATAC ATGGAGACAC ACCTGGGGCC ACCGCTGAGG GGCTCTACTG	360
GACCCCTAAT GGTGCGCGCC TGCCCTCTGA GCTGTCCCGC CTCCTTAACA CCTCCACCCT	420
GGCCCTGGCC CTGGCTAACC TTAATGGGTC CAGGCAGCAG TCAGGAGACA ATCTGGTGTG	480
TCACGCCCCG GACGGCAGCA TTCTGGCTGG CTCCTGCCTC TATGTTGGCT TGCCCCCTGA	540
GAAGCCCTTT AACATCAGCT GCTGGTCCCG GAACATGAAG GATCTCACGT GCCGCTGGAC	600
ACCGGGTGCA CACGGGGAGA CATTCTTACA TACCAACTAC TCCCTCAAGT ACAAGCTGAG	660
GTGGTACGGT CAGGATAACA CATGTGAGGA GTACCACACT GTGGGCCCTC ACTCATGCCA	720
TATCCCCAAG GACCTGGCCC TCTTCACTCC CTATGAGATC TGGGTGGAAG CCACCAATCG	780
CCTAGGCTCA GCAAGATCTG ATGTCTCAC ACTGGATGTC CTGGACGTGG TGACCACGGA	840
CCCCCACCC GACGTGCACG TGAGCCGCGT TGGGGGCCTG GAGGACCAGC TGAGTGTGCG	900
CTGGGTCTCA CCACCAGCTC TCAAGGATTT CCTCTTCCAA GCCAAGTACC AGATCCGCTA	960
CCGCGTGGAG GACAGCGTGG ACTGGAAGGT GGTGGATGAC GTCAGCAACC AGACCTCCTG	1020
CCGTCTCGCG GGCTGAAGC CCGGCACCGT TTACTTCGTC CAAGTGCCTT GTAACCCATT	1080
CGGGATCTAT GGGTCGAAA AGGCGGGAAT CTGGAGCGAG TGGAGCCACC CCACCGCTGC	1140
CTCCACCCCT CGAAGTGAGC GCCCGGGCCC GGGCGGCGGG GTGTGCGAGC CGCGGGGCGG	1200
CGAGCCCAGC TCGGGCCCCG TGCGGCGCGA GCTCAAGCAG TTCCTCGGCT GGCTCAAGAA	1260
GCAACGATAC TGCTCGAACC TTAGTTTCCG CCGTACGAC CAGTGGCCTG CTTGGATGCA	1320
GAAGTCACAC AAGACCCGAA ACCAGGTAGG AAGTTGGGG GAGGCTTGCG TGGGGGGTAA	1380
AGGAGCAGAG GAAGAGAGAG ACCCGGGTGA GCAGCCTCCA CAACACCGCA CTCTCTTTC	1440

09037657-031096

- 82 -

CAAGCACAGG ACGAGGGGAT CCTGCCCTCG GGCAGACGGG GTGCGGCGAG AGGTAAGGGG	1500
GTCTGGGTGA GTGGGGCCTA CAGCAGTCTA GATGAGGCC TTTCCCTCC TTCGGTGTG	1560
CTCAAAGGGA TCTCTTAGTG CTCATTTAC CCACTGCAAA GAGCCCCAGG TTTTACTGCA	1620
TCATCAAGTT GCTGAAGGGT CCAGGCTTAA TGTGGCCTCT TTTCTGCCCT CAGGTCCTGC	1680
CGGCTAAACT CTAAGGATAG GCCATCCTCC TGCTGGGTCA GACCTGGAGG CTCACCTGAA	1740
TTGGAGCCCC TCTGTACCTA TCTGGGCAAC AAAGAAACCT ACCATGAGGC TGGGGCACAA	1800
TGAGCTCCCA CAACCACAGC TTTGGTCCAC ATGATGGTCA CACTTGGATA TACCCAGTG	1860
TGGGTAAGGT TGGGGTATTG CAGGGCCTCC CAACAATCTC TTTAAATAAA TAAAGGAGTT	1920
GTTCAGGTAA	1930

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 560 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCCAGGCAGC GGTGCGGGGA CAACCTCGTG TGCCACGCCC GTGACGGCAG CATCCTGGCT	60
GGCTCCTGCC TCTATGTTGG CCTGCCCCCA GAGAAACCCG TCAACATCAG CTGCTGGTCC	120
AAGAACATGA AGGACTTGAC CTGCCGCTGG ACGCCAGGGG CCCACGGGGA GACCTTCCTC	180
CACACCAACT ACTCCCTCAA GTACAAGCTT AGGTGGTATG GCCAGGACAA CACATGTGAG	240
GAGTACCACA CAGTGGGGCC CCACTCCTGC CACATCCCCA AGGACCTGGC TCTCTTTACG	300
CCCTATGAGA TCTGGGTGGA GGCCACCAAC CGCCTGGGCT CTGCCCCTC CGATGTACTC	360
ACGCTGGATA TCCTGGATGT GGTGACCACG GACCCCCCGC CCGACGTGCA CGTGAGCCGC	420
GTCGGGGGCC TGGAGGACCA GCTGAGCGTG CGCTGGGTGT CGCCACCCGC CCTCAAGGAT	480
TTCCTTTTTC AAGCCAAATA CCAGATCCGC TACCGAGTGG AGGACAGTGT GGAATGGAAG	540
GTGGTGGACG ATGTGAGCAA	560

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1391 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1053

09037657.03109B

- 83 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ACC CTC AAC GGG CGC CGC CTG CCC CCT GAG CTC TCC CGT GTA CTC AAC	48
Thr Leu Asn Gly Arg Arg Leu Pro Pro Glu Leu Ser Arg Val Leu Asn	
1 5 10 15	
GCC TCC ACC TTG GCT CTG GCC CTG GCC AAC CTC AAT GGG TCC AGG CAG	96
Ala Ser Thr Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln	
20 25 30	
CGG TCG GGG GAC AAC CTC GTG TGC CAC GCC CGT GAC GGC AGC ATC CTG	144
Arg Ser Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu	
35 40 45	
GCT GGC TCC TGC CTC TAT GTT GGC CTG CCC CCA GAG AAA CCC GTC AAC	192
Ala Gly Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Val Asn	
50 55 60	
ATC AGC TGC TGG TCC AAG AAC ATG AAG GAC TTG ACC TGC CGC TGG ACG	240
Ile Ser Cys Trp Ser Lys Asn Met Lys Asp Leu Thr Cys Arg Trp Thr	
65 70 75 80	
CCA GGG GCC CAC GGG GAG ACC TTC CTC CAC ACC AAC TAC TCC CTC AAG	288
Pro Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys	
85 90 95	
TAC AAG CTT AGG TGG TAT GGC CAG GAC AAC ACA TGT GAG GAG TAC CAC	336
Tyr Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His	
100 105 110	
ACA GTG GGG CCC CAC TCC TGC CAC ATC CCC AAG GAC CTG GCT CTC TTT	384
Thr Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe	
115 120 125	
ACG CCC TAT GAG ATC TGG GTG GAG GCC ACC AAC CGC CTG GGC TCT GCC	432
Thr Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala	
130 135 140	
CGC TCC GAT GTA CTC ACG CTG GAT ATC CTG GAT GTG GTG ACC ACG GAC	480
Arg Ser Asp Val Leu Thr Leu Asp Ile Leu Asp Val Val Thr Thr Asp	
145 150 155 160	
CCC CCG CCC GAC GTG CAC GTG AGC CGC GTC GGG GGC CTG GAG GAC CAG	528
Pro Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln	
165 170 175	
CTG AGC GTG CGC TGG GTG TCG CCA CCC GCC CTC AAG GAT TTC CTC TTT	576
Leu Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe	
180 185 190	
CAA GCC AAA TAC CAG ATC CGC TAC CGA GTG GAG GAC AGT GTG GAC TGG	624
Gln Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp	
195 200 205	
AAG GTG GTG GAC GAT GTG AGC AAC CAG ACC TCC TGC CGC CTG GCC GGC	672
Lys Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly	
210 215 220	
CTG AAA CCC GGC ACC GTG TAC TTC GTG CAA GTG CGC TGC AAC CCC TTT	720
Leu Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe	
225 230 235 240	
GGC ATC TAT GGC TCC AAG AAA GCC GGG ATC TGG AGT GAG TGG AGC CAC	768
Gly Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His	
245 250 255	
CCC ACA GCC GCC TCC ACT CCC CGC AGT GAG CGC CCG GGC CCG GGC GGC	816
Pro Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly	
260 265 270	
GGG GCG TGC GAA CCG CGG GGC GGA GAG CCG AGC TCG GGG CCG GTG CGG	864
Gly Ala Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg	
275 280 285	

03037657-031099

- 84 -

CGC GAG CTC AAG CAG TTC CTG GGC TGG CTC AAG AAG CAC GCG TAC TGC	912
Arg Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys	
290 295 300	
TCC AAC CTC AGC TTC CGC CTC TAC GAC CAG TGG CGA GCC TGG ATG CAG	960
Ser Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln	
305 310 315 320	
AAG TCG CAC AAG ACC CGC AAC CAG CAC AGG ACG AGG GGA TCC TGC CCT	1008
Lys Ser His Lys Thr Arg Asn Gln His Arg Thr Arg Gly Ser Cys Pro	
325 330 335	
CGG GCA GAC GGG GCA CGG CGA GAG GTC CTG CCA GAT AAG CTG TAGGGGCTCA	1060
Arg Ala Asp Gly Ala Arg Arg Glu Val Leu Pro Asp Lys Leu	
340 345 350	
GGCCACCCTC CCTGCCACGT GGAGACGCAG AGGCCGAACC CAAACTGGGG CCACCTCTGT	1120
ACCCCTCACTT CAGGGCACCT GAGCCCTCA GCAGGAGCTG GGGTGGCCCC TGAGCTCCAA	1180
CGGCCATAAC AGCTCTGACT CCCACGTGAG GCCACCTTTG GGTGCACCCC AGTGGGTGTG	1240
TGTGTGTGTG TGAGGGTTGG TTGAGTTGCC TAGAACCCCT GCCAGGGCTG GGGGTGAGAA	1300
GGGGAGTCAT TACTCCCCAT TACCTAGGGC CCCTCCAAAA GAGTCCTTTT AAATAAATGA	1360
GCTATTTAGG TGCAAAAAA AAAAAAAAAA A	1391

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Thr Leu Asn Gly Arg Arg Leu Pro Pro Glu Leu Ser Arg Val Leu Asn	
1 5 10 15	
Ala Ser Thr Leu Ala Leu Ala Leu Asn Leu Asn Gly Ser Arg Gln	
20 25 30	
Arg Ser Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu	
35 40 45	
Ala Gly Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Val Asn	
50 55 60	
Ile Ser Cys Trp Ser Lys Asn Met Lys Asp Leu Thr Cys Arg Trp Thr	
65 70 75 80	
Pro Gly Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys	
85 90 95	
Tyr Lys Leu Arg Trp Tyr Gly Gln Asp Asn Thr Cys Glu Glu Tyr His	
100 105 110	
Thr Val Gly Pro His Ser Cys His Ile Pro Lys Asp Leu Ala Leu Phe	
115 120 125	
Thr Pro Tyr Glu Ile Trp Val Glu Ala Thr Asn Arg Leu Gly Ser Ala	
130 135 140	
Arg Ser Asp Val Leu Thr Leu Asp Ile Leu Asp Val Val Thr Thr Asp	
145 150 155 160	
Pro Pro Pro Asp Val His Val Ser Arg Val Gly Gly Leu Glu Asp Gln	
165 170 175	
Leu Ser Val Arg Trp Val Ser Pro Pro Ala Leu Lys Asp Phe Leu Phe	

09037657-031098

- 85 -

180	185	190
Gln Ala Lys Tyr Gln Ile Arg Tyr Arg Val Glu Asp Ser Val Asp Trp		
195	200	205
Lys Val Val Asp Asp Val Ser Asn Gln Thr Ser Cys Arg Leu Ala Gly		
210	215	220
Leu Lys Pro Gly Thr Val Tyr Phe Val Gln Val Arg Cys Asn Pro Phe		
225	230	235
Gly Ile Tyr Gly Ser Lys Lys Ala Gly Ile Trp Ser Glu Trp Ser His		
245	250	255
Pro Thr Ala Ala Ser Thr Pro Arg Ser Glu Arg Pro Gly Pro Gly Gly		
260	265	270
Gly Ala Cys Glu Pro Arg Gly Gly Glu Pro Ser Ser Gly Pro Val Arg		
275	280	285
Arg Glu Leu Lys Gln Phe Leu Gly Trp Leu Lys Lys His Ala Tyr Cys		
290	295	300
Ser Asn Leu Ser Phe Arg Leu Tyr Asp Gln Trp Arg Ala Trp Met Gln		
305	310	315
Lys Ser His Lys Thr Arg Asn Gln His Arg Thr Arg Gly Ser Cys Pro		
325	330	335
Arg Ala Asp Gly Ala Arg Arg Glu Val Leu Pro Asp Lys Leu		
340	345	350

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCCAGGCAGC GGTCCGGGGA CAAC

24

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TTGCTCACAT CGTCCACCAC CTTC

24

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6663 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

09037657-031098

- 86 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CCCAGAACTC TTGGACGCTG AGGCAGGAGG ATTCCCAAGT TTCAAGACAG TGTGTTTCTA	60
GGTAATGAGA CCCTGTCAAG AAAAGAAAAG AAATAAAGAG ACAAGAAAAT GTTTATAGGC	120
TGTGAGACAG CTTGGTGGGT AAGGGGCACT TGCCTCCAAT CAAGATGACC TCAGCCCCAT	180
CCCTAGGAAT CCATGOTAGA AGGAGAAAGC AAACCTCGAG CTGCTGACCT CCATACATGT	240
GCTCCAATGT GCACACACAC AGGGAGACAT AATCAATTAA TAGGATGTAT TTGCTTAGAT	300
TTGAGTAGGC ATTTATGACT GATGTTTTAA AATTTTATT TGATTTTATG AAAATATACC	360
TGTTTGTAT TGGTTTGTT TGGTTTGAGT TTTGTTTATT TGAGACAGG CTTCTCTGTG	420
TAGTCCTGGC TGTCTTGA ACTCACTCTG TAGACCAGG TGGCCTTGAA CTCAGAAATC	480
CGCCTGCTTG TGCTTCCAA GTGCTTAGAT TAAAGGTGTG CACTGCCATT CAGCAAAATT	540
GCATACTTA ACCCCAGTAT TTGGGAGGCA GAGGCAGACT AATGTGTGAA TTCCAGGCTA	600
GCCAAGGATA CAGAGTGAGA CCCTATTCTT ACCCTCCCC CCCAAAACCC CAAATGTAT	660
TTTGTGCTTG TGTATGTACA TGTGTGTTGC AGCACGTAAA TGTTCAAGGA CAACTGTAG	720
AAGTTCTCTC CGTTCACAGT CTAAGTCTG AATCAAAC AAGTCTCTA GGCTTAGCCA	780
CAGTCTTCTT TATGTACTGA GCCATTTCAC TGGCCCTGGA TTGACTGATG AATTAATTTT	840
TGAGATAAGG TCTCTGTAG CTCTAGCTAG GCTCAAACTA TGAACCTCCA AGGTCATCTT	900
GAGCTGCTGG TACTCTTGCT TCCACCCCAA GTGGTGAAT GATACTCAGG CAGCACTTCT	960
CTGGGGAAGG GGCTGGCCTT GGCCTTGATT TTGTTGCCTC AGCTTCAATG AGTGCTTGGG	1020
TCTCGTTGTT TCTTTTCTT ATCTGTGAAA TGGGTGAACA CCTGTTCAAG ACTTCCTGAC	1080
TCTTGAAACA TCCAGGCAGG GTGAGGACT TGAAGTGGGC TCATCCCATG CCTAACAAG	1140
TGTCGTCTTT GACCCAGAC ACAGCTGTAA TCAGCCCCCA GGACCCACC CTTCTCATCG	1200
GCTCCTCCCT GCAAGCTACC TGCTCTATAC ATGGAGACAC ACCTGGGGCC ACCGCTGAGG	1260
GGCTCTACTG GACCTCAAT GGTGCGCGCC TGGCCTCTGA GCTGTCCCGC CTCCTTAACA	1320
CCTCCACCCT GGCCCTGGCC CTGGCTAACC TTAATGGGTC CAGGCAGCAG TCAGGAGACA	1380
ATCTGGTGTG TCACGCCCGA GACGGCAGCA TTCTGGCTGG CTCCTGCCTC TATGTTGGCT	1440
GTAAGTGGGG CCCGAGAC TCAGAGATAG ATGGGGGTG GCAATGACAG ATTTAGAGCC	1500
TGGGTCTTCT GTCTGGGGC AGAGCCATGG GCTCTCACTT GCATGCAGGC ATGGTCATAC	1560
CCAGCACAGG CATTGCAACT CTAGGGACAG CTGTGGCTGC ACTGTCCCTT GTGTACCCCA	1620
CAGCTTTAGA AAAGCTGTCA TGTPTTCTT GTAGTGCCCC CTGAGAAGCC CTTTAACATC	1680
AGCTGCTGGT CCCGAACAT GAAGGATCTC ACGTCCGCT GGACACCGGG TGCACACGGG	1740
GAGACATTCT TACATACCAA CTACTCCCTC AAGTACAAGC TGAGGTTGGT ACCCAGCCAA	1800
GCCTTGCTGT GTGACTCTG GCAATACTTA CCTTCTCTGA TCAAATATGT TCCTGTTTAT	1860
GAACCAAAA GGGACTCTCG CACCTCCACA GGTGGTACGG TCAGGATAAC ACATGTGAGG	1920
AGTACCACAC TGTGGGCCCT CACTCATGCC ATATCCCAA GGACCTGGCC CTCTTCACTC	1980
CCTATGAGAT CTGGGTGGAA GCCACCAATC GCCTAGGCTC AGCAAGATCT GATGTCCTCA	2040

09037657 031098

- 87 -

CACTGGATGT CCTGGACGTG GGTGAGCCCC CAGTGTCCAC CTGTGTTCTG CCCTAGACCT 2100
 TATAGGGCGC CTCCCCCCCC TCCCCCAGA CTTTTTGGTT CTCTAGAGG TCTTAGCCAC 2160
 AGCCACGGTG GTTGCAGGAC AGTGGTTGTT CATAACTTAA TGCAAAGACT TTCCCCCAAG 2220
 ACAGTCAAGA TTTTTCCTT CCCCACCCCC AACACACACA TACACACACA CTCTGCAGAG 2280
 AACACCTGGC CTGACCACCC TCCCTCTCTA CAGCCCAGGT GTTCAGAAGG GAGTCCTAGG 2340
 GGACTGAGAG GAGGCGCCCA GGTCTGAAGG CGCCCCAGGA AGCCGAGGCC TTGAGCTGGG 2400
 GGGGGGGGCG AGGGTTGGAG GCACGAACTG GATGATCCCT GAGCACAACT GGGCCTAATC 2460
 TAATTAGGTT GTTCCCAGCC CAAAGCAGCC TGGGCCATTT AACCCCTCAA GTGCCTCACT 2520
 GAAGACTCAG GGGAGAGATC AGCTTGTACT CTCTCCATGG TCCCCCAGGA GGGTTCCTGG 2580
 GTGCCCTTGG CTCATTCCCC CATCCAGAGG TTTTGTGTCT TCCTGGCATC TAACCCTCAG 2640
 TTGTGCTCTG TGGCTGGCAC AGCTGCCCCG TGGAGGCTCT TGGTAATGTA CAAGGCATCA 2700
 GAGGTGGACA TGGGATGGGG ATACATAGGG ATGGAGCCAA ATAGCACCTC AAGGTGGGGT 2760
 GATATACAAT AAAGCTTGTC ACCCTGACGC TCAGAAAGCC TACTCATGAT GATCACAATT 2820
 GTTGACATCA CTCTGGGACA TGTAAGTGA CCCTAGCTCA AAACACAGAC AGTAGCTTTA 2880
 AGAGTCAGCT TGTGACTTAA TACTGGAAGT CAGGGCCTAA TAGGTGCTGG GTGATGCTCG 2940
 CCTCACTCCC TGTTTAGTGA GATCTCTGCG CTAATCTCCA CCCCAGCTGG GTGGGCTGCT 3000
 CTGTCCCCTT GAGGGCAGGA ATGTGTGTCT TCCATCAGAG ATAGGACCCG TGGTAGCAGC 3060
 AACTGCTGCT GGCTGTTTCT GGAATATTAA ATGACAGTAA TCTATCAGGC CTGGGTGAGT 3120
 AGCTAACAGG GGTGGGGGCG TGGTCTGGAA AACGCAGATA GGSTCATAGG AGCCACTGCA 3180
 GCCTAGATTA CACCACTGGG TGTTCTGTCA CTAGGCCATT CTCACCAAGC AGTCCTCAGA 3240
 ACTGGGAGCA CTGTTGCCAG CATTTAATGC CAGCATTTAA TGCCAGCATT AGGGGAGGCA 3300
 GAGGCAGAAG GATCTCTCTG AGTTCAAGGC CATCCTGAAT TTACATAAAG AGCTCCAGGC 3360
 CAGCCAGGGT GCGCAGTAAA ACCTTGCTCT AAAAACAATA GCATCTTTAG TGACCAGGCT 3420
 TGCTCCACCC CCAGTGACCA CGGACCCCCC ACCCGACGTG CACGTGAGCC GCGTTGGGGG 3480
 CCTGGAGGAC CAGCTGAGTG TGCGCTGGGT CTCACCACCA GCTCTCAAGG ATTCCTCTTT 3540
 CCAAGCCAAG TACCAGATCC GCTACCGCGT GGAGGACAGC GTGGACTGGA AGGTGCCCGT 3600
 CCGCCCCCGG ACCCGCCCCCT GACCCCGCCC CCGCATCTG ACTCCTCCCT CACCGTGCAG 3660
 GTGGTGGATG ACCTCAGCAA CCAGACCTCC TGCCGTCTCG CGGGCCTGAA GCCCCGCACC 3720
 GTTTACTTCG TCCAAGTGCG TTGTAACCCA TTCGGGATCT ATGGGTCGAA AAAGGCGGGA 3780
 ATCTGGAGCG AGTGGAGCCA CCCCACCGCT GCCTCCACCC CTCGAAGTGG TGAGCACCTC 3840
 TCCAGGGCTG GCTGGCCCAT GGAATCCCCA ATCCATCCTG TTCCTTCCCC CCCACCTTT 3900
 TTTTGAGACA GCGTCTTCAG GTAGCGCATG CTGGCCTTAA ATTCAGTATG TAGTCAAGGA 3960
 TGACCTCGAG CTCCTGGTCT TTTTGTCTCC ACTTAGAGAC AATGGCCAGT GGCCATCACC 4020
 ACCTTTGGGA GACTAGCCAT GGAGTCTATT TAGCCTGTCA TTTGGTGACA GATGGAGTAC 4080
 AACAGTGTGA CCTCTTGTA GAGAACTGAA GACAGGCTGT TTTTAACCCC AATATCCTAG 4140
 GCTCTCTAGA GGTAACTTT ATATAAATA GAGACTATTA CAGCCAGTTA TCACATGGTC 4200
 CCACAGAACC TTTTGTCA CAACCTATAG ACCACAGTGC CTGTGCCTAC CACATAAGGG 4260

09037657-031096

- 88 -

TCTCTACTGC	TGGCCACCC	CTCCAACCCT	TAAAAGGTAA	CCTAGGCAGC	CTTAATATTT	4320
GCAATCCTCC	TACCTCAGCC	TCTTGAATGC	TCAGAAACCA	GGCATTAAAC	CAAGTTTCTC	4380
TTCTCTGGGT	CCCTTTCTTA	AGGTGGGAGG	GCCTAAAGAT	GACTTCCTTT	GTCCTGAAGA	4440
CTCTCCGAGC	CCATGGATCT	GCACTCTCTA	ATATGAAATA	TATTGCATAA	AATGTCTGGC	4500
CTCAGTTTCC	CCACCTGTCA	GOTTTAGGCA	GCACAGTCGG	TCCAAGACAC	TTCATTATTT	4560
GCAGGCAGTA	TAAGAAGAAG	CTCCCATCCC	CCACCCGCTT	CCTCCGGTCC	CTAAGACAGA	4620
ATACTTCTAC	ACTGAAACTG	AACTCTCGCA	GACGCATATG	CTCACTTTAA	TGATGATGAA	4680
ATAATGGGGA	AACTGAGGCT	CCGAGAGATT	CCTGGAGGAA	GAGGGTCAAA	ACCAGCTCCA	4740
GGAGCTCTC	CAGCCCCCAT	CCGGGCTCT	CCAGGTTCG	GGCTTGCGG	GAGTGAACAC	4800
AGCTGGGAGG	GGCTGGAGCC	TGGGAGCTTT	GGCCCTTGCT	CGTGCCAGC	ACCTGCGATT	4860
CTTGACGGG	AGCCAGCAGG	CGGCTGCGTC	CGCCCGAGAG	ACTGAAGAAG	CCGGGGGTAG	4920
GGTTGGAGGG	AGGTAAGCAG	GGGCTGTGGG	GGCCGAAGCT	TGTGCCAGGG	CCTGTCAGCG	4980
AGTCCCCAGT	TTTATTTATG	GCGTGAGGCC	GATGTCCTTA	TCCGCTGGCC	TGCTGGGGGA	5040
TGGCTGCGGC	TGGGGATTGG	ACCCAAGGGC	TGGCTTCCCA	CTCAGTCCTC	CAGCCCACTC	5100
CATGTCACAC	CCGTGCATTC	TCTGAGGCTT	ATCTTGGGAA	CCCGCCCTTG	TTCTGTGCTG	5160
TCTGTCTCTA	TTTCTGTCTA	TCACTTTCCC	AGAGCCTTTT	TTTTATGCTT	TTAATATAAC	5220
TACGTTTTAA	AAATTGCTTT	TGTATAATGT	GTGTGCCTTC	GTGAGCGTGC	GTGCCACAAC	5280
ACACACGTGA	AGGTTAGAGA	ACTTTGTGTA	GTAGGCTCCT	TCCACCATGT	GGGACTAGGG	5340
CTGGCGACAA	GAGCAATTAC	TGAGTCATCT	CGCCAGCCCC	TCACCCCTCA	CTTCCCATCC	5400
TGTTTGGATA	GTCATAGGTA	ATCGAAGGTA	AATCGCTGGC	TTTAATTTTC	TAGCTATCCT	5460
GCCTCAGCCT	ACCAAGTGCT	GTGCTACCAC	GTTTGTGGGA	GGGGCTCTCC	TCCCAGTGTC	5520
TGGGGGTGAC	ACAGTCCCAA	GATCTCTGCT	TTCTAGGTCT	TTGTCTTAGT	TTGCCCCCTG	5580
CTTTGTCCGT	GTCCCTAGAG	TCTCCGGCCC	CACTTATCCA	TTGACTGGTC	TTTCCTTTAC	5640
CGAATACTCG	GTTTTACCTC	CCACTGATTT	GACTCCCTCC	TTTGCTTGTC	TCCATCGCCG	5700
TGGCATTGCC	ATTCTCTCTG	GTGACTCTGG	GTCCACACCT	GACACCTTTC	CCAACCTTCC	5760
CCAGCCGAAG	CTGGTCTGGT	ATGGGAGGCC	GCCGTCCCCG	GCGCGCTCC	TGCTGGCCGC	5820
GCCCCAACAC	TGCCGCTCCA	TTCTCTTAG	AGCGCCCGGG	CCCGGGCGGC	GGGGTGTGCG	5880
AGCCGCGGGG	CGGCGAGCCC	AGCTCGGGCC	CGGTGCGGCG	CGAGCTCAAG	CAGTTCCTCG	5940
GCTGGCTCAA	GAAGCACGCA	TACTGCTCGA	ACCTTAGTTT	CCGCCTGTAC	GACCAGTGGC	6000
GTGCTTGGAT	GCAGAAGTCA	CACAAGACCC	GAAACCAGGT	AGGAAAGTTG	GGGGAGGCTT	6060
GCGTGGGGGG	TAAAGGAGCA	GAGGAAGAGA	GAGACCCGGG	TGAGCAGCCT	CCACAACACC	6120
GCACTCTTCT	TTCCAAGCAC	AGGACGAGGG	GATCCTGCCC	TGGGCGAGAC	GGGGTGCGGC	6180
GAGAGGTAAG	GGGGTCTGGG	TGAGTGGGGC	CTACAGCAGT	CTAGATGAGG	CCCTTTCCCC	6240
TCCTTCGGTG	TTGCTCAAAG	GGATCTCTTA	GTGCTCATTT	CACCCACTGC	AAAGAGCCCC	6300
AGGTTTTACT	GCATCATCAA	GTTGCTGAAG	GGTCCAGGCT	TAATGTGGCC	TCTTTTCTGC	6360
CCTCAGGTCC	TGCCGGCTAA	ACTCTAAGGA	TAGGCCATCC	TCCTGCTGGG	TCAGACCTGG	6420
AGGCTCACCT	GAATTGGAGC	CCCTCTGTAC	CATCTGGGCA	ACAAAGAAAC	CTACCAGAGG	6480

09037657 031096

- 89 -

CTGGGCACAA TGAGCTCCCA CAACCACAGC TTTGGTCCAC ATGATGGTCA CACTTGGATA 6540
TACCCCAAGTG TGGGTAGGGT TGGGTATTG CAGGGCCTCC CAAGAGTCTC TTAAATAAA 6600
TAAAGGAGTT GTTCAGGTCC CGATGGCCAG TGTGTTTGGG GCCTATGTGC TGGGGTGGGG 6660
GGA 6663

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 186 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Asp Pro Thr Leu Leu Ile Gly Ser Ser Leu Gln Ala Thr Cys Ser Ile
1 5 10 15
His Gly Asp Thr Pro Gly Ala Thr Ala Glu Gly Leu Tyr Trp Thr Phe
20 25 30
Asn Gly Arg Arg Leu Pro Ser Glu Leu Ser Arg Leu Leu Asn Thr Ser
35 40 45
Thr Leu Ala Leu Ala Leu Ala Asn Leu Asn Gly Ser Arg Gln Gln Ser
50 55 60
Gly Asp Asn Leu Val Cys His Ala Arg Asp Gly Ser Ile Leu Ala Gly
65 70 75 80
Ser Cys Leu Tyr Val Gly Leu Pro Pro Glu Lys Pro Phe Asn Ile Ser
85 90 95
Cys Trp Ser Arg Asn Met Lys Asp Leu Thr Cys Arg Trp Thr Pro Gly
100 105 110
Ala His Gly Glu Thr Phe Leu His Thr Asn Tyr Ser Leu Lys Tyr Lys
115 120 125
Leu Arg Leu Val Arg Ser Gly * His Met * Gly Val Pro His Cys
130 135 140
Gly Pro Ser Leu Met Pro Tyr Pro Gln Gly Pro Gly Pro Leu His Ser
145 150 155 160
Leu * Asp Leu Gly Gly Ser His Gln Ser Pro Arg Leu Ser Lys Ile
165 170 175
* Cys Pro His Thr Gly Cys Pro Gly Arg
180 185

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

09037657-031098

- 90 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
AGCTGGCGCG CCTCCCGGGC GGATCGGGAG CCCAC

35

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AGCTACGCGT TTAGAGTTTA GCCGGCAG

28

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met	Val	Leu	Ala	Ser	Ser	Thr	Thr	Ser	Ile	His	Thr	Met	Leu	Leu	Leu
1				5				10					15		
Leu	Leu	Met	Leu	Phe	His	Leu	Gly	Leu	Gln	Ala	Ser	Ile	Ser		
		20					25					30			

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ile	Lys	Pro	Ser	Gly	Arg	Arg	Gly	Ala	Ala	Arg	Gly	Pro	Ala	Gly	Asp	Tyr	Lys	Asp	Asp
					5				10					15				20	
Asp	Asp	Lys																	

(2) INFORMATION FOR SEQ ID NO:34:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 73 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

00037657-031098

- 91 -

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATCTTGCCC TCGGGCAGAC GGGGTGCGGC GAGAGGTCTT GCCGGCGACT ACAAGGACGA
 CGATGACAAG TAG

60

73

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 73 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AACGGGAGCC CGTCTGCCCC ACGCCGCTCT CCAGGACGGC CGCTGATGTT CCTGCTGCTA
 CTGTTCATCC TAG

60

73

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CCCACGCTTC TCATCGGATT CTCCTG

27

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CAGTCCACAC TGTCTCCAC TCGGTAG

27

(2) INFORMATION FOR SEQ ID NO:38:

09037657-031096

- 92 -

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11832 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GCGGCCGCTG	CAGTGATTAC	TCACCGCGTG	GCGCACCCCA	CCCGCGGGCC	GCTGAGTGA	60
TTTTTCCGTG	GGGGGATGTG	AAGAAGTTTA	GGGAGAACTC	TTCTGCACCG	ATGGGAACATA	120
GGAATGCAGG	GTTCCGTTCC	GTTCCCCAAA	GGACACACCT	CTCCCCATAA	GCCCACTCAT	180
AAGGGCTCCC	TGCACGCGCT	CCGGGACATC	CCCATATCCA	ATACCCGCAG	ATATGATAGT	240
TGAGAAGGGA	CCAGAGGCCG	GAGACTCCCT	CCCTGCCTTC	TGGCTTTCCC	CCCCCCCCTG	300
ACGAAACGAG	ACTACAGCGA	TGGGAGAGGT	GGCATGAAGG	CTTAGGGTGG	GGATCGGTAG	360
GACCCATGCA	CCCAGAGAAA	GGGACTGGTG	GCAACTTTCA	AACTCTCTGG	GGAAGGAAGA	420
AGGGCTGAAA	GAGGATGAAC	GGGCTCAGGT	ACTGCTCAAT	GTGTGTGTGG	CGGACCAAAG	480
TGGGTATGGG	GGCCCCGTAA	GAGGGGCGGG	GAAGTGGAT	AGGAAGGATC	CCGGTAGACT	540
GGAGGGGATC	CTGAAAAGC	ACCAGGGCTG	CGAGCTAGGA	ACCCATTCCG	AGTTAAGGGT	600
ACAGGATCCC	AGATGAGGGG	GTGGGAAGCC	TGGGACGGGC	GGGACCAGAG	AGGGAGGTCC	660
CACGGGCTGG	TGGGGAAGA	GTGGGGGGCT	TGCGCGAGGA	GGATGGGACG	TTCAGGAGTG	720
GTAAGTGGGC	GGAGGCCCGC	CGGGCGGGGC	GCGCGGTGCC	CGCGGGCGGT	GGGAAGGCCG	780
GTGCGGGGCC	CACGATCAAC	CCCCCCCCAG	GGGCCGGGCC	GGGCCGGGGG	CGGGGCCGGG	840
CGGGGCGAGC	GGCGATTAG	CGCCTTGTC	ATTTCGGCTG	CTCAGACTTG	CTCCGGCCTT	900
CGCTGTCCGC	GCCCAGTGAC	GCGCGTGAGG	ACCCGAGCCC	CAATCTGCAC	CCCGCAGACT	960
CGCCCCCGCC	CCATACCGGC	GTTGCAGTCA	CGGCCCGTTG	CGCGCCACCC	CCATGCCCGC	1020
GGGTGCGCCC	GGCCCCGTGC	CCCAATCCGC	GCGCGGGCCG	CCGCGGCCGC	TGTCCTCGCT	1080
GTGGTGCCTT	CTGTTGCTCT	GTGCTCTCG	GGTGCTCTCG	GGCGGATCGG	GAGCCCGTGA	1140
GTACCGTGCG	CCCTGCTCCC	CACCTCCCCA	GGGAAGCCGG	GATCCGGCGC	CCCGGGGGGT	1200
AGTCGCGGGG	GATGGAAGAA	GGGGCGCGAG	CGCCACCTGG	ACGTCCCGGG	AACAAAGGAA	1260
GGCGGCCCTC	GGGGCGCCCT	CACCTGTGGG	GCTCATGGCA	CCACCACCCA	GCCTCCCAAG	1320
AGTACCCCGT	TATACATCAG	AGGCCTCTTA	TCTGTATCCC	CTTTGCGAGG	CTGTCTGGCC	1380
AGGCTCAGTT	TGAAGGACAT	CGCAGTGTC	TGGGACCCCC	CTCCTTCAGG	GTGCTGGGAC	1440
GCTTCGGGGC	GCACGCCTGT	GTCTTGATA	TCAGAGCGGA	AGGGAAGCCT	CCCTGGCCGG	1500
GGGCGCACGC	TTGGGTGCGT	TGGGTTGGGT	GCTGGCGCAA	AGTGGGGTCC	CCTCCCCCAT	1560
GAAGTGATGA	TCCCCGGGGG	GAGGGTGGGG	CGTTATCGTG	AGCCCTCCTG	TCCGCTGGC	1620
ATGCGGCCCC	GCCTCCCTCG	GGACTTGCCT	CTCCGTGGGG	TGGCGCCCGC	CCCCTCCCCC	1680
CTATAGCAGA	CTCCATGCTT	TGGTATCCTC	GAAGTCTCTT	CCACTGGTGG	GGCTCACAAC	1740
CGGTCTCATT	CAGGCTGCGC	TGGGTTGAGA	GCCTCTAGCG	ACTGAAATTT	CGGTGAGGAG	1800
CGAGAGCAAG	CGTGTCCGGG	CACCGCGAGC	CCAGACTTCA	TTGTCTAAGG	GGCACCACAGT	1860

09037657 "031008

- 93 -

GGGGGTCAGC	TGCCGAGAGA	ATCCCACTGT	CCCAGGAGGA	ACTCCTGGCC	TTGAGCCCCC	1920
ATCACCCAAC	GCACACATCC	CCGCCAGGAT	GCGGTCTCCA	CATCCAGACC	CTCTCTGGGA	1980
CACACCCAAA	GACACACAAA	AGAGCCCCAC	TGGCTTATGT	CCCGTCACCC	TGCCCTCCGA	2040
CGCGCGCTGC	AGCCCAGATG	CGTATTGCA	CACCATCGCG	GCGCTCGCAT	TCCATCCTCT	2100
ACACACACAC	ACACACACAC	ACACACACAC	ACACACACAC	ACACACAGAC	ACGCACACAC	2160
ACACGCACGC	ACACACACGC	ACGCCCCGAC	TCGTGGTCCC	ACATTTATTT	CACAGGGGAG	2220
GCAACACCGG	GGTACGCATA	TGGTTGAGTG	CACTGGAGAT	CTTTCCCCAC	CACTCTCAGG	2280
ACCCCATCCG	GAGACACAGG	CCACACCGCA	GGGGCACCAC	GCTGCGCTGC	TGCTCTGGGC	2340
TAGTAGTCTT	GTGCAGTTTG	TCCGCGGTGT	CTGTGGACGC	CCTCCCGCTC	TTGTCAGGGG	2400
ACAGGAACCT	ACACTCCTGC	TTGCCCAAGG	CGGCTGGGCA	GGTGATGTGG	TGACACCCGG	2460
GACCTTTCCG	GGGAGTTGGT	GTTGCTGCCA	AGCCTGGGTA	GTTTTTGAAT	GCCACCAATA	2520
GCGCTAAGCT	TTGTTTCCGG	GCGGGCTGCA	GAGCAACAGG	CGAAGGTGGC	GGAGTGGGGG	2580
TGGCGCGTGT	GTTTTTTCTT	TTAAGGGGGA	GAGAAATTAA	ATAAGAGGTT	CTCACACCTC	2640
TGCAATCTGT	TTGTACTTAC	CGTGTGTCTT	AACACCTGAC	CAGCCAGCCG	GTGGGTCTGT	2700
AAAGTGATG	CAGGTACCAG	CGGGACAGGA	GATGGGGGCC	CCTGGGGTAT	GGCTGGGATG	2760
GAGGCCACCT	TCCCGTTGGC	CTTTCAGGGA	ATCTCACACT	TTTCCCTTTT	AAAACACATG	2820
GTGTTCTTTT	TAATAACGGC	AGCAACTCCG	CATTGGGAAA	GGGGGAAATA	AGCTTGTATA	2880
GGCCCCGGCT	TTGTGGAAAG	GAGGGGAAGA	GGGAAGAAAA	AAGGAGGGGT	GTCTCCTCCA	2940
GGCTTAGGGG	GCTGTACAGT	GCTGCTCTGT	CTAGCTTGGC	ATGTGTGTGC	CCCAGTCCCC	3000
AGTGGCTTTG	GCCCATGTGT	TGTGGAAGCC	AAGAGGGAGA	CTGGAGTCCT	CTATCTCTGG	3060
TACTCCAGAG	TCAGGCTTCT	CAGTCCGAGC	CCAGAGAACG	TCTTCCCTGT	TTTATGGAGG	3120
GAATCAGGGA	AGGGGGTGCC	AGGTGGACTA	CGTTCTGCTG	AGGACTGTAC	CAGTCGCTCG	3180
AAGGAGAAAG	CTTGGGCTTG	CCCCCTCCC	CCCTCAAGCC	ACGAAGGGCA	GCTGCTAGGC	3240
TAGTGTGGTA	AAAGGGCATT	ACTCCCCAGC	CAGGACCCCC	CAGAGAGTCC	CCTTCCTGGC	3300
CAGACAAATG	CTGGGGAGGG	ACAGAGGGGT	GTGATCATTG	CCCAGGAGTG	CAGACAGTGG	3360
GGTCCCGGGT	CGGGCAGTGC	CTCCACCCCT	GCTGAGGGGG	GCGCCAGGC	AGGAAGCGGT	3420
GGGTGGGCGG	GGGTAGAGAC	GCTGGCACGT	CCCAGTTCAT	GCCGAAGGAA	TTCTGAATTA	3480
GCGGGCGGCT	GGCTGCCTGG	GACCTCCGGG	GCGGCCCCCT	GGCCCCCGCC	GCTCCGTCTG	3540
GCCTGCTCCT	CCTGCTCCTT	CGCACGGACG	CTGAGACCTC	CGCTGAGCCC	TGGGACAAGC	3600
CCCAAATGCA	ACTGCGATTG	CAGGCTTCGC	AAGACCCGCC	TCCTCCCAAG	GCCAAATTTG	3660
CCTGGGAGAA	GTCATTTCAGG	GCCCAGACTA	GAACCATGTT	GGTGCCACCT	CATCCATCTG	3720
GGGCATGAAG	GACCGTCCAG	GGCTGCAGTT	TAGCTTCTTA	ATAGGAACCT	GGGGGTGGGT	3780
GCAGCCTCTG	TTCTCCGAGC	CTCTTTGGAA	ATCGGTTTGT	TTTTTGTTTT	TGTTTTTTCC	3840
AATACTCTTT	TCCTCTCATC	CCATCCCCGG	ACTGTTTTCC	TCCCTAAGGG	TTGAGAGCCC	3900
TGCAGTCTTC	CCTAACCTTT	TCTTTGCTTC	TACCCACGGG	CCTTTGCACA	TGGAGTCCCA	3960
CCTCTCCCCCT	TGCCCAACTG	GGGCTCCAGC	CTTACTGCAT	TTGGCTCTTG	GTAAGTGTCC	4020
CAGGGCCTCT	CTGACACACA	GGGTTGTAGC	CCCAGCTCCC	TCTCTTCTCC	TCCCCCCTTT	4080

05037657-031096

CTCTTTTGCT TCTGAGACTT AATTTTTC TTTTCTTTT TGGCTTTTG AGACAGGGTT	4140
TCTCTGTACA GCCCTGGCTG CCCTGGCACT CATTCTGTAG ACCAGGCTAG CCTCAAACCTC	4200
ACAAACCTAC CTGCTCTGC CTTTCCAGTG CTGGCACTAA AGATGTGGGC CACCACAAC	4260
AGTAGTTAAG TGTTTGCTG TGTCTTATT CCTATAGTGA CCTCAGTTC TGGCATATTG	4320
TAGGCGATGG ATGGATGAAT GGATGGATGG ATGGATGGAT GGATGGTTGG ATGGAGCAAG	4380
CTTGAATCOT CCTGAGTGAA AAAAGAGACC TCAGAGAAGT GAATGGAGTT AGGTTCCAG	4440
GGCAGCCTGG CCTGCTGGTC TCATGGGAGC TCCCTGTGAA ACTTCCCCCA CACCTCCAC	4500
CACCTTGCCA TCCTGTGTGG CTGACAAGAA AGGCCAATGG CCAGATGGGG ACACAGACTC	4560
AGGGAAGCTT GGAATATGTT CCCCTCCTCA TATCCTAGGC CTTGTGTGCC CCCTGAGGGC	4620
CCAGCCTATG AGTAGGGCAG CTGTGGGCTG CCCTAAGGTT GGGTAGGCAA GAAGGGGGTG	4680
GTCCCTCAGG GTGGGTCACA GGATGAGGT CATTTCAAA GTGGCCATCA CAGTGCCCT	4740
AGGAAATGAT TGTGGAGAGT CAGAACTCCT GTTGGGAGTT GTAGAGGGCC TTGCATGTGG	4800
GCTTCTGTGG CTGTCCCTTC TCTGTGGTC CTTTGCACAG TCCCTCTGTG TGTGCTGGGA	4860
TGTGAGGAGG GCACGGGGAA AATGAAGGCT CAGCCCTCA GCTTCCCTT CACGOTTCAC	4920
CCAACAGGGC TCACCTCTCC TCTGGACAGG CTCTCACTGT ATGCACAGAT TGGCCTCACA	4980
TTTGATTCCC TTCCTTTGGT CTCTGGGAT GACAAACATT TACCAGGGTA GGATTTTACA	5040
TTTAGATAT GTCCATTCTC CAGAAACACA CTTGTGAGGT TAGGGTATCA GTGAAAGGAC	5100
ACCACCAGGA CAGACAAAGA ATTGGAGAGG AAGGAAATG GTAAGCCAGG CCATGCTTGA	5160
TGGCTTATGT GTAATCCAG AACTCTGGAC GCTGAGGCAG GAGGATTCCA AGTTTCAAGA	5220
CAGTGTGTTT TAGGTAATGA GACCCTGTCA AGAAAAGAAA AGAAATAAAG AGACAAGAAA	5280
ATGTTTATAG GCTGTGAGAC AGCTTGGTGG GTAAGGGGCA CTTGCCTCCA ATCAAGATGA	5340
CCTCAGCCCC ATCCCTAGGA ATCCATGGTA GAAGGAGAAA GCAAACCTCA GCTGCTGACC	5400
TCCATACATG TGCTCCAATG TGACACACA CAGGGAGACA TAATCAATTA ATAGGATGTA	5460
TTTGCTTAGA TTTGAGTAGG CATTTATGAC TGATGTTTAA AAATTTTAT TTGATTTAT	5520
GAAAATATAC CTGTTTGTAT TTGGTTTGGT TTGGTTTGG TTTTGTATTT TTGAGACAGG	5580
GCTTCTCTGT GTAGTCTGG CTGTCTTGG AACTCACTCT GTAGACCAGG CTGGCCTTGA	5640
ACTCAGAAAT CCGCTGCTT GTGCTTCCCA AGTGCTTAGA TTAAAGGTGT GCACTGCCAT	5700
TCAGCAAAAT TGCATACTTT AACCCAGTA TTTGGGAGGC AGAGGCAGAC TAATGTGTGA	5760
ATTCCAGGCT AGCCAAGGAT ACAGAGTGAG ACCCTATTCT TACCCTCCCC CCCCAAAACC	5820
CCAAAATGTA TTTTGTGCTT GTGTATGTAC ATGTGTGTG CAGCACGTAA ATGTCCAAGG	5880
ACAACCTGTA GAAGTTCTCT CCGTTCACAG TCTAAGTCCT GAATTCAAAC TAAGGTCCTC	5940
AGGCTTAGCC ACAGTCTTCT TTATGTACTG AGCCATTTC CTGGCCCTGG ATGACTGAT	6000
GAATTAATTT TTGAGATAAG GTCTCTTGTA GCTCTAGCTA GGCTCAAAC ATGAACCTCC	6060
AAGGTCATCT TGAGCTGCTG GTACTCTTGC TTCCACCCCA AGTGGTGGA TGATACTCAG	6120
GCAGCACTTC TCTGGGAAG GGGCTGGCCT TGGCCTTGAT TTTGTTGCCT CAGCTTCAAT	6180
GAGTGCTTGG GTCTCGTTGT TTCTTTTCTT TATCTGTGAA ATGGGTGAAC ACCTGTTCAA	6240
GACTTCTGA CTCTTGAAAC ATCCAGGCAG GGTGAGGGAC TTGAAGTGG CTCATCCCAT	6300

09037657 "031096

GCCTAACAAA GTGTCGTCTT TGACCCCA GACAGCTGTA ATCAGCCCCC AGGACCCAC 6360
 CCTTCTCATC GGCTCCTCCC TGCAAGCTAC CTGCTCTATA CATGGAGACA CACCTGGGGC 6420
 CACCGCTGAG GGGCTCTACT GGACCTTCAA TGGTCGCCGC CTGCCCTCTG AGCTGTCCCG 6480
 CCTCCTTAAC ACCTCCACCC TGGCCCTGGC CCTGGCTAAC CTTAATGGGT CCAGGCAGCA 6540
 GTCAGGAGAC AATCTGGTGT GTCACGCCGC AGACGGCAGC ATTCTGGCTG GCTCCTGCCT 6600
 CTATGTTGGC TGTAAGTGGG GCCCCAGACA CTCAGAGATA GATGGGGGTT GGCAATGACA 6660
 GATTTAGAGC CTGGGTCTTC TGTCTGGGG CAGAGCCATG GGCTCTCACT TGCATGCAGG 6720
 CATGGTCATA CCCAGCACAG GCATTGCAAC TCTAGGGACA GCTGTGGCTG CACTGTCCCC 6780
 TGTGTACCCC ACAGCTTTAG AAAAGCTGTC ATGTTTTCTT TGTAGTGCCC CCTGAGAAGC 6840
 CCTTTAACA CAGCTGCTGG TCCCGGAACA TGAAGGATCT CACGTGCCGC TGGACACCGG 6900
 GTGCACACGG GGAGACATTC TTACATACCA ACTACTCCCT CAAGTACAAG CTGAGGTTGG 6960
 TACCCAGCCA AGCCTTGCTG TGTGACTTCT GGCAATACTT ACCTTCTCTG ATCAAATATG 7020
 TTCCTGTTTA TGAAGTCAA AGGGACTCTC GCACCTCCAC AGGTGGTACG GTCAGGATAA 7080
 CACATGTGAG GAGTACCACA CTGTGGGCCC TCACTCATGC CATATCCCCA AGGACCTGGC 7140
 CCTCTTCACT CCCTATGAGA TCTGGGTGGA AGCCACCAAT CGCCTAGGCT CAGCAAGATC 7200
 TGATGTCTC AACTGGATG TCCTGGACGT GGGTGAGCCC CCAGTGCCA CCTGTGTTCT 7260
 GCCCTAGACC TTATAGGGCG CCTCCCCCCC ATCCCCCAG ACTTTTGGT TCTTCTAGAG 7320
 GTCTTAGCCA CAGCCACGGT GGTTCAGGA CAGTGGTTGT TCATAACTTA ATGCAAAGAC 7380
 TTTCCCCAA GACAGTCAAG ATTTTCCCT CCCCACCCC AACACACACA TACACACACA 7440
 CTCTGCAGAG AACACCTGGC CTGACCACCC TCCCTCTCTA CAGCCCAGGT GTTCAGAAGG 7500
 GAGTCCTAGG GGAAGTGAAG GAGGCGCCCA GGTCTGAAGG CGCCCCAGGA AGCCGAGGCC 7560
 TTGAGCTGGG GGGGGGGCG AGGGTTGGAG GCACGAACTG GATGATCCCT GAGCACAAC 7620
 GGGCCTAATC TAATTAGGGT GTTCCAGCC CAAAGCAGCC TGGGCCATT AACCCTTCAA 7680
 GTGCCTCACT GAAGACTCAG GGGAGAGATC AGCTTGTA CTCTCCATGG TCCCCAGGA 7740
 GGGTTCCTGG GTGCCCCTGG CTCATTCCCA CATCCAGAGG TTTGTGTCT TCCTGGCATC 7800
 TAACCCTCAG TTGTGCTCTG TGGCTGGCAC AGCTGCCCCG TGGAGGCTCT TGGTAATGTA 7860
 CAAGGCATCA GAGGTGGACA TGGGATGGG ATACATAGG ATGGAGCCAA ATAGCACCTC 7920
 AAGGTGGGGT GATATACAAT AAAGCTTGTC ACCCTGACGC TCAGAAAGCC TACTCATGAT 7980
 GATCACAATT GTTGACATCA CTCTGGGACA TGTAGTGAGA CCCTAGCTCA AAACACAGAC 8040
 AGTAGCTTTA AGAGTCAGCT TGTGACTTAA TACTGGAAC CAGGGCCTAA TAGGTGCTGG 8100
 GTGATGCTCG CCTCACTCCC TGTTTAGTGA GATCTCTGCG CTAATCTCCA CCCCAGCTGG 8160
 GTGGGCTGCT CTGTCCCCTT GAGGGCAGGA ATGTGTGTCT TCCATCAGAG ATAGGACCCG 8220
 TGGTAGCAGC AACTGCTGCT GGCTGTTTCT GGAATATTAA ATGACAGTAA TCTATCAGGC 8280
 CTGGGTGAGT AGCTAACAGG GGTGGGGCG TGGTCTGGAA AACGCAGATA GGGTCATAGG 8340
 AGCCACTGCA GCCTAGATTA CACCACTGGG TGTTCTGTCA CTAGGCCATT CTCACCAAGC 8400
 AGTCCTCAGA ACTGGGAGCA CTGTTGCCAG CATTTAATGC CAGCATTTAA TGCCAGCATT 8460
 AGGGGAGGCA GAGGCAGAAG GATCTCTCTG AGTTCAAGGC CATCCTGAAT TTACATAAAG 8520

09037657.031096

AGCTCCAGGC	CAGCCAGGGT	GCGCAGTAAA	ACCTTGTCTC	AAAAAACAAA	GCATCTTTAG	8580
TGACCAGGCT	TGCTCCACCC	CCAGTGACCA	CGGACCCCCC	ACCCGACGTG	CACGTGAGCC	8640
GCGTGGGGG	CCTGGAGGAC	CAGCTGAGTG	TGCGCTGGGT	CTCACCACCA	GCTCTCAAGG	8700
ATTTCTCTT	CCAAGCCAAG	TACCAGATCC	GCTACCGCGT	GGAGGACAGC	GTGGACTGGA	8760
AGGTGCCCCG	CCCGCCCCCG	ACCCGCCCTT	GACCCCGCCC	CCCGCATCTG	ACTCCTCCCT	8820
CACCGTGACG	GTGGTGGATG	ACGTCAGCAA	CCAGACCTCC	TGCCGTCTCG	CGGGCCTGAA	8880
GCCCGGCACC	GTTTACTTCG	TCCAAGTGCG	TTGTAACCCA	TTGCGGATCT	ATGGGTGCGA	8940
AAAGGCGGGA	ATCTGGAGCG	AGTGGAGCCA	CCCCACCGCT	GCCTCCACCC	CTCGAAGTGG	9000
TGAGCACCTC	TCCAGGGCTG	GCTGGCCCAT	GGAATCCCCA	ATCCATCCTG	TTCTTCCCCC	9060
CCCACCCCTT	TTTTGAGACA	GCGTCTTCAG	GTAGCGCATG	CTGGCCTTAA	ATTCACTATG	9120
TAGTCAAGGA	TGACCTCGAG	CTCCTGGTCT	TTTTGTCTCC	ACTTAGAGAC	AATGGCCAGT	9180
GGCCATCACC	ACCTTTGGGA	GACTAGCCAT	GGAGTCTATT	TAGCCTGTCA	TTTGGTGACA	9240
GATGGAGTAC	AACAGTGTGA	CCTCTTGTA	GAGAACTGAA	GACAGGCTGT	TTTTAACCCC	9300
AATATCCTAG	GCTCTCTAGA	GGTTAACTTT	ATATAAAATA	GAGACTATTA	CAGCCAGTTA	9360
TCACATGGTC	CCACAGAACC	TTTTGTGACA	CAACCTATAG	ACCACAGTGC	CTGTGCCTAC	9420
CACATAAGGG	TCTCTACTGC	TGGCCCCACC	CTCCAACCCT	TAAAAGGTAA	CCTAGGCAGC	9480
CTTAATATTT	GCAATCCTCC	TACCTCAGCC	TCTTGAATGC	TCAGAAACCA	GGCATTAAAC	9540
CAAGTTTCTC	TTCTCTGGGT	CCCTTTCTTA	AGGTGGGAGG	GCCTAAAGAT	GACTTCCTTT	9600
GTCTGAAGA	CTCTCCGAGC	CCATGGATCT	GCACTCTCTA	ATATGAAATA	TATTGCATAA	9660
AATGTCTGGC	CTCAGTTTCC	CCACCTGTCA	GGTTTAGGCA	GCACAGTCGG	TCCAAGACAC	9720
TTCAATATTT	GCAGGCAGTA	TAAGAAGAAG	CTCCCATCCC	CCACCCGCTT	CCTCCGGTCC	9780
CTAAGACAGA	ATACTTCTAC	ACTGAACTG	AACTCTCGCA	GACGCATATG	CTCACTTTAA	9840
TGATGATGAA	ATAATGGGGA	AACTGAGGCT	CCGAGAGATT	CCTGGAGGAA	GAGGGTCAAA	9900
ACCAGCTCCA	GGAAGCTCTC	CAGCCCCCAT	CCGGGCTCTT	CCAGGTTCTG	GGCTTGGCGG	9960
GAGTGAACAC	AGCTGGGAGG	GGCTGGAGCC	TGGGAGCTTT	GGCCCTTGCT	CGTCCCCAGC	10020
ACCTGCGATT	CTTGACGGG	AGCCAGCAGG	CGGCTGCGTC	CGCCCGAGAG	ACTGAAGAAG	10080
CCGGGGGTAG	GGTTGGAGGG	AGGTAAGCAG	GGGCTGTGGG	GGCCGAAGCT	TGTGCCAGGG	10140
CCTGTCAGCG	AGTCCCCAGT	TTTATTTATG	GCGTGAGGCC	GATGTCCTTA	TCCGCTGGCC	10200
TGCTGGGGGA	TGGCTGCGGC	TGGGATTGG	ACCCAAGGGC	TGGCTTCCCA	CTCAGTCCTC	10260
CAGCCCACTC	CATGTCACAC	CCGTGCATTC	TCTGAGGCTT	ATCTTGGGAA	CCCGCCCTTG	10320
TTCTGTGCTG	TCTGTCTCTA	TTTCTGTCAT	TCACTTTCCC	AGAGCCTTTT	TTTTATGCTT	10380
TTAATATAAC	TACGTTTTAA	AAATTGCTTT	TGTATAATGT	GTGTGCCTTC	GTGAGCGTGC	10440
GTGCCACAAC	ACACACGTGA	AGGTTAGAGA	ACTTTGTTGA	GTAGGCTCCT	TCCACCATGT	10500
GGGACTAGGG	CTGGCGACAA	GAGCAATTAC	TGAGTCATCT	CGCCAGCCCC	TCACCCCTCA	10560
CTTCCCATCC	TGTTTGGATA	GTCATAGGTA	ATCGAAGGTA	AATCGCTGGC	TTTAATTTTCG	10620
TAGCTATCCT	GCCTCAGCCT	ACCAAGTGCT	GTGCTACCAC	GTTTGTGGGA	GGGGCTCTCC	10680
TCCCAGTGTC	TGGGGGTACA	CAGTCCCAAG	ATCTCTGCTT	TCTAGGTCTT	TGTCTTAGTT	10740

09037657 031008

- 97 -

TGCCCCCTTGC TTTGTCCGTG TCCCTAGAGT CTCCGGCCCC ACTTAGTCTC CATTGATTTC	10800
CTTTCTGACC GAATACTCGG TTTTACCTCC CACTGATTG ACTCCCTCCT TTGCTTGTCT	10860
CCATCGCCCT GGCATTGCCA TTCCTCTGGG TGA CTCTGGG TCCACACCTG ACACCTTTCC	10920
CAACTTTCCC CAGCCGAAGC TGGTCTGGTA TGGGAGGCCG CGTCCCCGCG CGCGCCTCCT	10980
GCTGGCCGCG CCCCCAACT GCCGCTCCAT TCTCTTTAGA GCGCCCGGGC CCGGGCGGCG	11040
GGGTGTGCGA GCCGCGGGGC GCGGAGCCCA GCTCGGGCCC GGTGCGGCGC GAGCTCAAGC	11100
AGTTCTCTCG CTGGCTCAAG AAGCAGCAT ACTGCTCGAA CCTTAGTTTC CGCCTGTACG	11160
ACCACTGGCG TGCTTGGATG CAGAAGTCAC ACAAGACCCG AAACCAGGTA GGAAAGTTGG	11220
GGGAGGCTTG CGTGGGGGGT AAAGGAGCAG AGGAAGAGAG AGACCCGGGT GAGCAGCCTC	11280
CACAACACCG CACTCTTCTT TCCAAGCACA GGACGAGGGG ATCCTGCCCT CGGGCAGACG	11340
GGGTGCGGCG AGAGGTAAGG GGGTCTGGGT GAGTGGGGCC TACAGCAGTC TAGATGAGGC	11400
CCTTTCCCCCT CCTTCGGTGT TGCTCAAAGG GATCTCTTAG TGCTCATTTT ACCCACTGCA	11460
AAGAGCCCCA GGTTTTACTG CATCATCAAG TTGCTGAAGG GTCCAGGCTT AATGTGGCCT	11520
CTTTTCTGCC CTCAGGTCCT GCCGGCTAAA CTCTAAGGAT AGGCCATCCT CCTGCTGGGT	11580
CAGACCTGGA GGCTCACCTG AATTGGAGCC CCTCTGTACC ATCTGGGCAA CAAAGAAACC	11640
TACCAGAGGC TGGGCACAAAT GAGCTCCAC AACCACAGCT TTGGTCCACA TGATGGTCAC	11700
ACTTGGATAT ACCCCAGTGT GGGTAGGGTT GGGGTATTGC AGGGCCTCCC AAGAGTCTCT	11760
TTAAATAAAT AAAGGAGTTG TTCAGGTCCC GATGGCCACT GTGTTTGGGG CCTATGTGCT	11820
GGGGTGGGGG GA	11832

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 amino acids
 - (B) TYPE: amino acids
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Val	Ile	Ser	Pro	Gln	Asp	Pro	Thr	Leu	Leu	Ile	Gly	Ser	Ser	Leu	Gln	Ala	Thr	Cys	Ser
			5						10					15					20
Ile	His	Gly	Asp	Thr	Pro														
				25															

09037657-031098

- 98 -

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GTCCAAGTGC GTTGTAAACC A

21

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCTGAGTGTG CGCTGGGTCT CACC

24

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GGCTCCACTC GCTCCAGA

18

05037657-031098

45
B